

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-338869

(43) 公開日 平成11年(1999)12月10日

(51) Int.Cl.<sup>6</sup>

識別記号

F I

G 0 6 F 17/30

G 0 6 F 15/403

3 4 0 A

3 2 0 D

審査請求 未請求 請求項の数18 O L (全 14 頁)

(21) 出願番号 特願平10-142864

(22) 出願日 平成10年(1998)5月25日

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 鷺崎 誠司

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

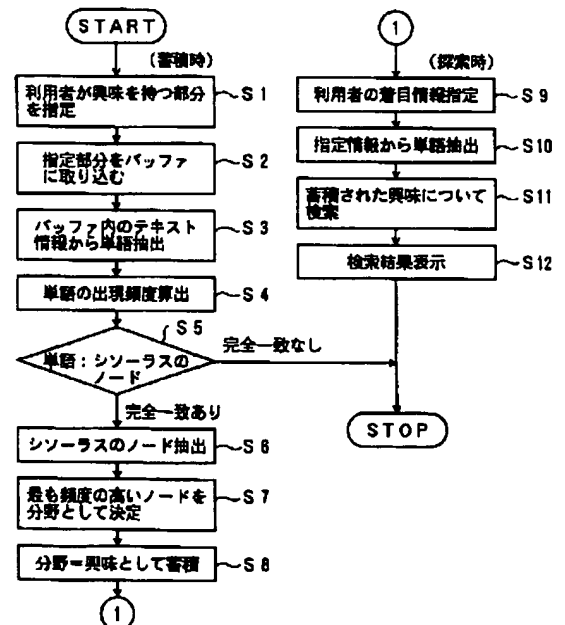
(54) 【発明の名称】 情報推薦方法及びシステム及び情報推薦プログラムを格納した記憶媒体及び情報蓄積方法及び装置及び情報蓄積プログラムを格納した記憶媒体

(57) 【要約】

【課題】 情報提供側で情報が関連づけられていない場合であっても、システムに問い合わせが可能な情報推薦方法及びシステム及び情報推薦プログラムを格納した記憶媒体及び情報蓄積方法及び装置及び情報蓄積プログラムを格納した記憶媒体を提供する。

【解決手段】 本発明は、参照情報に対して、利用者が興味を持つ部分を指定し、指定された部分をバッファに取り込み、取り込んだバッファ内のテキスト情報から単語を抽出し、抽出した単語の出現頻度を計算し、出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、当該ノードを抽出し、各々の単語に対して抽出したシソーラスの該当ノードから最も頻度の高いノードを、該情報が属する分野として決定して蓄積する。参照時は、利用者による指定された情報の参照時に着目した情報から単語を抽出し、抽出した単語に該当する情報を蓄積した興味から検索し、検索された結果を表示することで、他利用者の知識を利用する。

本発明の原理を説明するための図



## 【特許請求の範囲】

【請求項 1】 膨大な情報の中から利用者の目的と情報を探索する際に、他利用者が持つ情報を利用して情報探索を行うための情報推薦方法において、参照情報に対して、利用者が興味を持つ部分を指定し、指定された部分をバッファに取り込み、取り込んだ前記バッファ内のテキスト情報から単語を抽出し、抽出した単語の出現頻度を計算し、前記出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、当該ノードを抽出し、各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該情報が属する分野として決定し、利用者が選択した複数の情報に対して、決定した前記分野を、利用者毎の興味と見做し、該興味を利用者毎に蓄積し、前記利用者による情報の参照時に着目した情報を指定し、指定した情報から単語を抽出し、抽出した単語に該当する情報を蓄積した前記興味から検索し、検索された結果を表示することで、他利用者の知識を利用することを特徴とする情報推薦方法。

【請求項 2】 前記各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理する際に、前記各々の単語に対して抽出されるノードの重複を計算し、前記各ノードの頻度を算出する請求項 1 記載の情報推薦方法。

【請求項 3】 前記シソーラスに相当するものとして、予め多数の情報を階層的に分類し、各々の情報を抽象化する単語（タグ）を付与しておき、前記抽出した単語に対して、階層的に分類した情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する請求項 1 記載の情報推薦方法。

【請求項 4】 膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報推薦システムであって、参照情報に対して利用者が興味を持つ部分を指定する参照情報指定手段と、前記参照情報指定手段により指定された部分を記憶手段に取り込む指定部分格納手段と、前記記憶手段内のテキスト情報から単語を抽出する単語抽出手段と、前記単語抽出手段で抽出された単語の出現頻度を計算する単語出現頻度計算手段と、前記単語出現頻度計算手段で計算された出現頻度の高い

単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出する単語抽出手段と、各々の単語に対して前記単語抽出手段で抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該当情報が属する分野として決定する分野決定手段と、利用者が選択した複数の情報に対して、前記分野決定手段で決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段に格納する興味情報格納手段と、前記利用者による情報の参照時に着目した情報を指定する情報指定手段と、前記情報指定手段で指定された情報から単語を抽出する指定単語抽出手段と、前記指定単語抽出手段で抽出された単語に該当する情報を前記付加情報記憶手段に蓄積されている前記興味から検索する興味検索手段と、前記興味検索手段で検索された結果を表示する結果表示手段とを有することを特徴とする情報推薦システム。

【請求項 5】 前記分野決定手段は、前記同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理手段を含む請求項 4 記載の情報推薦システム。

【請求項 6】 前記分野決定手段は、前記シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類手段と、各々の情報を抽象化する単語（タグ）を付与するタグ付与手段と、前記単語抽出手段で抽出された単語に対して、前記階層分類手段で分類された情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する手段を含む請求項 5 記載の情報推薦システム。

【請求項 7】 膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報推薦プログラムを格納した記憶媒体であって、参照情報に対して利用者が興味を持つ部分を指定させる参照情報指定プロセスと、前記参照情報指定プロセスにより指定された部分を記憶手段に取り込む指定部分格納プロセスと、前記記憶手段内のテキスト情報から単語を抽出する単語抽出プロセスと、前記単語抽出プロセスで抽出された単語の出現頻度を計算する単語出現頻度計算プロセスと、前記単語出現頻度計算プロセスで計算された出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出する単語抽出プロセスと、

各々の単語に対して前記ノード抽出プロセスで抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該当情報が属する分野として決定する分野決定プロセスと、

利用者が選択した複数の情報に対して、前記分野決定プロセスで決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段に格納する興味情報格納プロセスと、

前記利用者による情報の参照時に着目した情報を指定させる情報指定プロセスと、

前記情報指定プロセスで指定された情報から単語を抽出する指定単語抽出プロセスと、

前記指定単語抽出プロセスで抽出された単語に該当する情報を前記付加情報記憶手段に蓄積されている前記興味から検索する興味検索プロセスと、

前記興味検索プロセスで検索された結果を表示させる結果表示プロセスとを有することを特徴とする情報推薦プログラムを格納した記憶媒体。

【請求項 8】 前記分野決定プロセスは、前記同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理プロセスを含む請求項 7 記載の情報推薦プログラムを格納した記憶媒体。

【請求項 9】 前記分野決定プロセスは、前記シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類プロセスと、各々の情報を抽象化する単語（タグ）を付与するタグ付与プロセスと、

前記単語抽出プロセスで抽出された単語に対して、前記階層分類プロセスで分類された情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定するプロセスを含む請求項 7 記載の情報推薦プログラムを格納した記憶媒体。

【請求項 10】 膨大な情報の中から利用者の目的と情報を探索する際に、他利用者が持つ情報を利用して情報探索を行うための情報を蓄積する情報蓄積方法において、

参照情報に対して、利用者が興味を持つ部分を指定し、指定された部分をバッファに取り込み、

取り込んだ前記バッファ内のテキスト情報から単語を抽出し、

抽出した単語の出現頻度を計算し、

前記出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、当該ノードを抽出し、

各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該情報が属する分野として決定し、

利用者が選択した複数の情報に対して、決定した前記分野を、利用者毎の興味と見做し、該興味を利用者毎に記

憶手段に蓄積することを特徴とする情報蓄積方法。

【請求項 11】 前記各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理する際に、前記各々の単語に対して抽出されるノードの重複を計算し、

前記各ノードの頻度を算出する請求項 10 記載の情報蓄積方法。

【請求項 12】 前記シソーラスに相当するものとして、予め多数の情報を階層的に分類し、各々の情報を抽象化する単語（タグ）を付与しておき、前記抽出した単語に対して、階層的に分類した情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する請求項 10 記載の情報蓄積方法。

【請求項 13】 膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報を蓄積する情報蓄積装置であって、

参照情報に対して利用者が興味を持つ部分を指定する参照情報指定手段と、

前記参照情報指定手段により指定された部分を記憶手段に取り込む指定部分格納手段と、

前記記憶手段内のテキスト情報から単語を抽出する単語抽出手段と、

前記単語抽出手段で抽出された単語の出現頻度を計算する単語出現頻度計算手段と、

前記単語出現頻度計算手段で計算された出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出するノード抽出手段と、

各々の単語に対して前記ノード抽出手段で抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該当情報が属する分野として決定する分野決定手段と、

付加情報を格納する付加情報記憶手段と、

利用者が選択した複数の情報に対して、前記分野決定手段で決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段に格納する興味情報格納手段とを有することを特徴とする情報蓄積装置。

【請求項 14】 前記分野決定手段は、前記同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理手段を含む請求項 13 記載の情報蓄積装置。

【請求項 15】 前記分野決定手段は、前記シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類手段と、各々の情報を抽象化する単語（タグ）を付与するタグ付与手段と、

前記単語抽出手段で抽出された単語に対して、前記階層分類手段で分類された情報に含まれる文字列と一致する

ものに付与されているタグを、該当情報が属する分野として決定する手段を含む請求項13記載の情報蓄積装置。

【請求項16】 膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報を蓄積する情報蓄積プログラムを格納した記憶媒体であって、参照情報に対して利用者が興味を持つ部分を指定させる参照情報指定プロセスと、前記参照情報指定プロセスにより指定された部分を記憶手段に取り込む指定部分格納プロセスと、前記記憶手段内のテキスト情報から単語を抽出する単語抽出プロセスと、前記単語抽出プロセスで抽出された単語の出現頻度を計算する単語出現頻度計算プロセスと、前記単語出現頻度計算プロセスで計算された出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出するノード抽出プロセスと、各々の単語に対して前記ノード抽出プロセスで抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該当情報が属する分野として決定する分野決定プロセスと、利用者が選択した複数の情報に対して、前記分野決定プロセスで決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段に格納する興味情報格納プロセスとを有することを特徴とする情報蓄積プログラムを格納した記憶媒体。

【請求項17】 前記分野決定プロセスは、前記同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理プロセスを含む請求項16記載の情報蓄積プログラムを格納した記憶媒体。

【請求項18】 前記分野決定プロセスは、前記シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類プロセスと、各々の情報を抽象化する単語（タグ）を付与するタグ付与プロセスと、前記単語抽出プロセスで抽出された単語に対して、前記階層分類プロセスで分類された情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定するプロセスを含む請求項16記載の情報蓄積プログラムを格納した記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報推薦方法及びシステム及び情報推薦プログラムを格納した記憶媒体及び情報蓄積方法及び装置及び情報蓄積プログラムを格納した記憶媒体に係り、特に、膨大な情報の中から自分の目的とするものを探索するための情報探索装置におけ

る、情報の探索のために、他者が持つ情報を利用して情報探索を効率良く行うための情報推薦方法及びシステム及び情報推薦プログラムを格納した記憶媒体及び情報蓄積方法及び装置及び情報蓄積プログラムを格納した記憶媒体に関する。

【0002】

【従来の技術】情報探索時における情報推薦方法では、予め関連付けられている情報に対して匿名利用者が何等かの方法で評価を下し、その評価に基づき適当な情報を適切な利用者に提供している。利用者が自分の評価を付与する情報の単位としては、文書全体または、予め付加情報が存在するという印を付けられた情報に限られている。そのような情報を様々な利用者が参照することで、該当利用者はその情報に対して興味を持っていると判断される。各利用者が興味を持った情報は蓄積されて、他の利用者の要求に従いアクセスされる。

【0003】この場合、目的とする情報にアクセスしたい利用者は、予め情報に対して静的に関連付け（例えば、ハイパーリンク情報が付与されている）がされている情報や、TITAN (<http://sting.navi.ntt.co.jp/>) などの検索エンジンと呼ばれる検索システムを用いて情報を探索する。TITANのような検索システムでは、予め多くの情報を手元に収集しておき、情報に含まれる単語の出現特性に基づいて、利用者の検索要求に最も合致する情報を表示する機能を持つ。利用者は疑問に感じたことを検索語句に変換して検索システムに問い合わせると、システムはそれに最も見合う情報を画面上に表示する利用者は表示された情報の中から適切なものを選択する。

【0004】

【発明が解決しようとする課題】しかしながら、上記従来のシステムは、情報提供側で情報が関連付けられていない場合には、システムに問い合わせを行うことができないという問題がある。さらに、推薦する情報の単位がページ単位であるので、利用者が不要な情報も混在し、利用者が必要な情報のみを抽出するために利用者による取捨選択が必要である。

【0005】本発明は、上記の点に鑑みなされたもので、情報提供側で情報が関連づけられていない場合であっても、システムに問い合わせが可能な情報推薦方法及びシステム及び情報推薦プログラムを格納した記憶媒体及び情報蓄積方法及び装置及び情報蓄積プログラムを格納した記憶媒体を提供することを目的とする。

【0006】

【課題を解決するための手段】図1は、本発明の原理を説明するための図である。本発明（請求項1）は、膨大な情報の中から利用者の目的と情報を探索する際に、他利用者が持つ情報を利用して情報探索を行うための情報推薦方法において、参照情報に対して、利用者が興味を持つ部分を指定し（ステップ1）、指定された部分をバ

ッファに取り込み（ステップ2）、取り込んだバッファ内のテキスト情報から単語を抽出し（ステップ3）、抽出した単語の出現頻度を計算し（ステップ4）、出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し（ステップ5）、完全一致するものが存在する場合、当該ノードを抽出し（ステップ6）、各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該情報が属する分野として決定し（ステップ7）、利用者が選択した複数の情報に対して、決定した分野を、利用者毎の興味と見做し、該興味を利用者毎に蓄積し（ステップ8）、利用者による情報の参照時に着目した情報を指定し（ステップ9）、指定した情報から単語を抽出し（ステップ10）、抽出した単語に該当する情報を蓄積した興味から検索し（ステップ11）、検索された結果を表示する（ステップ12）ことで、他利用者の知識を利用する。

【0007】本発明（請求項2）は、各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理する際に、各々の単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する。本発明（請求項3）は、シソーラスに相当するものとして、予め多数の情報を階層的に分類し、各々の情報を抽象化する単語（タグ）を付与しておき、抽出した単語に対して、階層的に分類した情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する。

【0008】図2は、本発明の原理構成図である。本発明（請求項4）は、膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報推薦システムであって、参照情報に対して利用者が興味を持つ部分を指定する参照情報指定手段110と、参照情報指定手段110により指定された部分を記憶手段220に取り込む指定部分格納手段210と、記憶手段220内のテキスト情報から単語を抽出する単語抽出手段230と、単語抽出手段230で抽出された単語の出現頻度を計算する単語出現頻度計算手段240と、単語出現頻度計算手段240で計算された出現頻度の高い単語に対して、予め保持しているシソーラス60中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出するノード抽出手段250と、各々の単語に対してノード抽出手段で抽出されたシソーラス60の該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該当情報が属する分野として決定する分野決定手段260と、利用者が選択した複数の情報に対して、分野決定手段260で決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段50に格納する興味情報格納手段270と、利用者による情報の参照時に着目した情報を指定する情報指定手段120と、情報指定手

段120で指定された情報から単語を抽出する指定単語抽出手段280と、指定単語抽出手段280で抽出された単語に該当する情報を付加情報記憶手段50に蓄積されている興味から検索する興味検索手段290と、興味検索手段290で検索された結果を表示する結果表示手段295とを有する。

【0009】本発明（請求項5）は、分野決定手段260において、同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理手段を含む。本発明（請求項6）は、分野決定手段260において、シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類手段と、各々の情報を抽象化する単語（タグ）を付与するタグ付与手段と、単語抽出手段で抽出された単語に対して、階層分類手段で分類された情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する手段を含む。

【0010】本発明（請求項7）は、膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報推薦プログラムを格納した記憶媒体であって、参照情報に対して利用者が興味を持つ部分を指定させる参照情報指定プロセスと、参照情報指定プロセスにより指定された部分を記憶手段に取り込む指定部分格納プロセスと、記憶手段内のテキスト情報から単語を抽出する単語抽出プロセスと、単語抽出プロセスで抽出された単語の出現頻度を計算する単語出現頻度計算プロセスと、単語出現頻度計算プロセスで計算された出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出するノード抽出プロセスと、各々の単語に対してノード抽出プロセスで抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該当情報が属する分野として決定する分野決定プロセスと、利用者が選択した複数の情報に対して、分野決定プロセスで決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段に格納する興味情報格納プロセスと、利用者による情報の参照時に着目した情報を指定させる情報指定プロセスと、情報指定プロセスで指定された情報から単語を抽出する指定単語抽出プロセスと、指定単語抽出プロセスで抽出された単語に該当する情報を付加情報記憶手段に蓄積されている興味から検索する興味検索プロセスと、興味検索プロセスで検索された結果を表示させる結果表示プロセスとを有する。

【0011】本発明（請求項8）は、分野決定プロセスにおいて、同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理プロセスを含む。本発明（請求項9）は、分野決定プロセスにおいて、シソーラスに相当するものとし

て、予め多数の情報を階層的に分離する階層分類プロセスと、各々の情報を抽象化する単語（タグ）を付与するタグ付与プロセスと、単語抽出プロセスで抽出された単語に対して、階層分類プロセスで分類された情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定するプロセスを含む。

【0012】本発明（請求項10）は、膨大な情報の中から利用者の目的と情報を探索する際に、他利用者が持つ情報を利用して情報探索を行うための情報を蓄積する情報蓄積方法において、参照情報に対して、利用者が興味を持つ部分を指定し、指定された部分をバッファに取り込み、取り込んだバッファ内のテキスト情報から単語を抽出し、抽出した単語の出現頻度を計算し、出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、当該ノードを抽出し、各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該情報が属する分野として決定し、利用者が選択した複数の情報に対して、決定した分野を、利用者毎の興味と見做し、該興味を利用者毎に記憶手段に蓄積する。

【0013】本発明（請求項11）は、各々の単語に対して抽出したシソーラスの該当ノードから、同一情報を整理する際に、各々の単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する。本発明（請求項12）は、シソーラスに相当するものとして、予め多数の情報を階層的に分類し、各々の情報を抽象化する単語（タグ）を付与しておき、抽出した単語に対して、階層的に分類した情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する。

【0014】本発明（請求項13）は、膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報を蓄積するための情報蓄積装置であって、参照情報に対して利用者が興味を持つ部分を指定する参照情報指定手段と、参照情報指定手段により指定された部分を記憶手段に取り込む指定部分格納手段と、記憶手段内のテキスト情報から単語を抽出する単語抽出手段と、単語抽出手段で抽出された単語の出現頻度を計算する単語出現頻度計算手段と、単語出現頻度計算手段で計算された出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出するノード抽出手段と、各々の単語に対してノード抽出手段で抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該情報が属する分野として決定する分野決定手段と、付加情報を格納する付加情報記憶手段と、利用者が選択した複数の情報に対して、分野決定手段で決定された該当分野を、利用者毎の興味と見做し、該興味を利

用者毎に付加情報記憶手段に格納する興味情報格納手段とを有する。

【0015】本発明（請求項14）は、分野決定手段において、同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理手段を含む。本発明（請求項15）は、分野決定手段において、シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類手段と、各々の情報を抽象化する単語（タグ）を付与するタグ付与手段と、単語抽出手段で抽出された単語に対して、階層分類手段で分類された情報に含まれる文字列と一致するものに付与されているタグを、該当情報が属する分野として決定する手段を含む。

【0016】本発明（請求項16）は、膨大な情報の中から利用者の目的とするものを探索する際に、他利用者が持つ情報を利用して、情報探索を行うための情報を蓄積するための情報蓄積プログラムを格納した記憶媒体であって、参照情報に対して利用者が興味を持つ部分を指定させる参照情報指定プロセスと、参照情報指定プロセスにより指定された部分を記憶手段に取り込む指定部分格納プロセスと、記憶手段内のテキスト情報から単語を抽出する単語抽出プロセスと、単語抽出プロセスで抽出された単語の出現頻度を計算する単語出現頻度計算プロセスと、単語出現頻度計算プロセスで計算された出現頻度の高い単語に対して、予め保持しているシソーラス中のノードと比較し、完全一致するものが存在する場合、該当ノードを抽出するノード抽出プロセスと、各々の単語に対してノード抽出プロセスで抽出されたシソーラスの該当ノードから、同一情報を整理した後で、最も頻度の高いノードを、該情報が属する分野として決定する分野決定プロセスと、利用者が選択した複数の情報に対して、分野決定プロセスで決定された該当分野を、利用者毎の興味と見做し、該興味を利用者毎に付加情報記憶手段に格納する興味情報格納プロセスとを有する。

【0017】本発明（請求項17）は、分野決定プロセスにおいて、同一情報を、各々野単語に対して抽出されるノードの重複を計算し、各ノードの頻度を算出する同一情報整理プロセスを含む。本発明（請求項18）は、分野決定プロセスにおいて、シソーラスに相当するものとして、予め多数の情報を階層的に分離する階層分類プロセスと、各々の情報を抽象化する単語（タグ）を付与するタグ付与プロセスと、単語抽出プロセスで抽出された単語に対して、階層分類プロセスで分類された情報に含まれる文字列と一致するものに付与されているタグを、該情報が属する分野として決定するプロセスを含む。

【0018】本発明は、情報閲覧中に、情報中にある部分的な文字列に着目して適当な情報の蓄積や推薦を行う。即ち、利用者は情報の蓄積時には、自分が興味のある部分はどこでもよいので、指定し、その部分が属する

分野を参照情報に付加して全体を自動的に蓄積する。こうして、多くの人により蓄積された情報を相互に検索可能にしておくことで、ある利用者の情報参照時に、陽に付加情報が付与されていない場合でも、調査したい部分を指定すれば、蓄積した情報へ自動的にアクセスできるため、効率のよい検索ができる。

【0019】特に有効なのは、ある分野に通じた利用者が多くの情報を蓄積している状態で、利用者が関連情報を自分で、あるいは、ハイパーリンク情報として簡単に探せない場合、システムに問い合わせることができるため、厳選された情報が容易に入手可能となる。さらに、従来では、推薦する情報の単位がページ単位などと大きく、その情報の中には様々な情報が混在していることが多い。その場合、本当に必要な情報だけではないので、利用者側で取捨選択の操作が必要となる。本発明によれば、他の利用者により情報中からその人が必要な部分だけを選択して蓄積されているため、蓄積された情報の中から類似した情報を抽出すれば、真に有効な情報の提供が可能となる。

【0020】

【発明の実施の形態】図3は、本発明の情報推薦システムの構成を示す。同図に示す情報推薦システムは、検索モジュール10と蓄積モジュール20、情報30、付加情報40、付加情報データベース50及びシソーラス格納部60から構成される。

【0021】検索モジュール10は、領域選択部11、情報検索部12及び情報表示部13より構成され、蓄積モジュール20を探索する。蓄積モジュール20は、領域選択部21、情報分類部22、及び情報蓄積部23から構成され、利用者が興味を持った情報を格納する。上記の処理を図4、図5のフローチャートを用いて説明する。

【0022】図4は、本発明の蓄積モジュールの動作を示すフローチャートである。まず、蓄積モジュール20における処理について説明する。

ステップ101) 利用者は領域選択部21において、参照した情報の中から自分が興味のある部分(領域)を指定する。領域に含まれる文字列がバッファ(図示せず)に格納される。

【0023】ステップ102) バッファに格納された文字列は、情報分類部22に渡され、情報分類部22では、文字列に含まれる単語(名詞)を一般的な形態素解析装置を用いて抽出する。

ステップ103) 抽出された単語の特徴に基づいて情報の分類を行う。これは、例えば、特願平10-45770で実現される情報分類情報を用いることで可能である。

【0024】ステップ104) ステップ102において、各々の文字列が含まれる領域が属する分野が決定されると、情報蓄積部23は、文字列そのものと、それが

含まれる分野を付加情報データベース50に蓄積する。情報の再利用が可能なように、各文字列は、単語が抽出され、単語の重みを利用して検索が可能なようにインデックスデータベース(図示せず)を作成しておき、URLと分野を関連付けて保持する。

【0025】図5は、本発明の検索モジュールの動作を示すフローチャートである。

ステップ201) 情報選択部11は、自分の探索目的に近い領域を選択する。

ステップ202) 情報検索部12は、情報選択部11で選択された領域に含まれる文字列から単語を抽出する。

【0026】ステップ203) 情報検索部12は、抽出された単語に関する分類先を決定する。

ステップ204) 当該分類先に関する付加情報データベース50に対して問い合わせを行い、類似する情報が当該付加情報データベース50に存在する場合には、ステップ205に移行し、存在しない場合にはステップ206に移行する。

【0027】ステップ205) 付加情報データベース50に蓄積されている情報の中から問い合わせに用いた単語に最も近いと判断された情報を、基の参照情報の指定した領域に付随させて表示する。

ステップ206) 付加情報データベース50に類似する情報が存在しない場合には、付加情報は抽出されなかったもので、情報のみを表示する。

【0028】

【実施例】以下、図面と共に本発明の実施例を説明する。以下の実施例では、利用者がWWW上の情報を参照している状況を考える。あるURL(Uniform Resource Locator)で表される情報の例を図6に示す。図6に示す情報の先頭には、見出しに相当するような文字が含まれることが多く、その下には写真などと並んで詳細情報が構成されている。図6の例では、全体がスポーツの情報であり、その中でもサッカーと野球に関する情報が含まれている。

【0029】利用者は、領域選択部21において、このように参照した情報の中から自分が興味のある部分を指定する。例えば、ある利用者がサッカーに興味があり、野球に興味がない場合は、サッカーに関する部分だけをマウス操作などにより領域指定して、バッファに取り込む。マウス操作による領域指定とバッファに取り込む方法に関しては、計算機上で動作しているウィンドウシステムに依存するが、例えば、UNIX上のX-Windowシステムの場合は、マウスの左クリックにより領域の角を指定して、そのままドラッグすることにより領域を作成する。最終的に望みの領域が囲むことができれば、マウスのボタンを離すことにより領域が確定する。領域が確定された時点で、領域に含まれる文字列は、X-Windowシステムが持つ情報操作バッファ中に取り込まれる(ス

テップ101)。

【0030】バッファに取り込まれた文字列は、情報分類部22に渡される。情報分類部22では、文字列に含まれる単語(名詞)を一般的な形態素解析装置を用いて抽出して(ステップ102)、その特徴を基に情報の分類を行う(ステップ103)。これは、例えば、特願平10-45770で実現される情報分類方法を用いることで可能である。

【0031】この分類の処理例を以下に示す。分類処理には、図7に示すシソーラス格納部60に格納されているシソーラスを用いる。一般的にシソーラスとは、様々な概念で表現されるものを、各々の概念間の上位・下位関係に沿って木構造で表現したものである。図7の例では、一番上記のROOTノードから、マルチメディア、スポーツなどのいくつかのノードに枝分かれし、更に、スポーツノードは、野球、サッカーなどのノードに細分化されている様子が分かる。本発明では、このような一般的なシソーラス体系を用いる手法以外に、各ノードに対して各々に該当するような情報が蓄積されているものも利用できる。即ち、図7に示すサッカーノード以下に、タイトルが「サッカー掲示板」、「ワールドカップ」などの情報が蓄積されているものである。後者の例として、WWW上の検索サービスである“Yahoo!”(<http://www.yahoo.co.jp/>)が持つ階層構造がある。まず、上記処理により情報から抽出した各単語に対してシソーラス中のノードに一致するものを探索する。こうして得られるノード情報を全単語に対して収集し、最も頻度の高いノードの情報を、各々の文字列が属する分野と見做すことにする。ノードに情報が蓄積されている場合は、予め蓄積されている情報から単語を抽出してデータベース化しておき、文字列から抽出した各単語に対して、シソーラス中に蓄積された情報のどれに相当するかを探索する。合致した情報が属するノードを、各々の文字列が属する分野と見做すことにする。

【0032】こうして、各々の文字列が含まれる領域が属する分野が決定されると、情報蓄積部23では、文字列そのものと、それが含まれる分野を付加情報データベース50に蓄積する。情報の再利用が可能のように、各文字列は単語が抽出され、単語の重みを利用して検索が可能のようにインデックスデータベース(図示せず)を作成しておき、URLと分野を関連付けて保持する。付加情報データベース50の情報保有の概念図を図8に示す。

【0033】図8では、ある情報から領域が選択されて、その領域がある分野に分類された上で、付加情報データベース50に蓄積されるところを図式化している。付加情報データベース50内では、利用者により切り取られた情報が効率よく、検索可能のように整理されて蓄積されている。この情報は、様々な利用者により蓄積されるため、多様な分野に関する情報が幅広く蓄積される

ことになる。この例では、利用者A～Cによる5つの情報を蓄積しており、各々は、スポーツ関係や音楽関係に分類されて蓄積されていることを表している。

【0034】こうして、付加情報データベース50に情報が蓄積されている状態で、ある利用者が自分の目的に沿って情報を探索している状況を考える。その場合、ある起点となるWWW上の情報30を、例えば、TITANなどのような探索システムで発見し、参照する。その情報は有効であるかもしれないし、そうでない可能性もある。もしも、自分の目的には大体合っているが、もう少し詳細な情報が欲しいような場合に、情報選択部11の機能により、自分の探索目的に近い領域を選択する。

【0035】例えば、その領域にはサッカーに関する事項が記載されているとすると、情報探索部12に、その選択された領域に含まれる文字列が渡され、単語が抽出されて付加情報データベース50に問い合わせを行う。すると、付加情報データベース50に蓄積されている情報の中から問い合わせに用いた単語に最も近いと判断された情報が抽出される。

【0036】こうして得られた情報は、情報表示部13で現在参照している情報30と共に利用者に提供される。図9の例では、指定した領域に含まれる文字列を検索語句として、付加情報データベース50を探索した際に、最も合致する情報を抽出して、元の参照情報30の指定した領域に付随させて表示することで、利用者に付加情報を提供する。

【0037】また、上記の実施例は、図3に示す構成に基づいて説明しているが、この例に限定されることなく、検索モジュール10の各構成要素及び蓄積モジュール20の各構成要素をプログラムとして構築し、情報推薦装置として利用されるコンピュータに接続されるディスクや、フロッピーディスクやCD-ROM等の可搬記憶媒体に格納しておき、本発明を実施する際にインストールすることにより、容易に本発明を実現することが可能である。

【0038】また、図3に示す蓄積モジュール20のみを情報蓄積装置として独立して構築することも可能であり、検索モジュール10については、既存の検索システムを適用して、情報蓄積装置内の付加情報データベース50を検索するような構成であってもよい。なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

【0039】

【発明の効果】上述のように、本発明によれば、情報の閲覧中に、情報中にある部分的な文字列に着目して、適当な情報の蓄積や推薦を行うため、情報提供側で情報が関連付けられていない場合でも、システムに問い合わせることができるため、厳選された情報が簡単に入手することができる。この方法を用いることにより、よりきめ細かな情報推薦サービスの構築が可能になる。



## 【図面の簡単な説明】

【図 1】 本発明の原理を説明するための図である。

【図 2】 本発明の原理構成図である。

【図 3】 本発明の情報推薦システムの構成図である。

【図 4】 本発明の蓄積モジュールの動作を示すフローチャートである。

【図 5】 本発明の検索モジュールの動作を示すフローチャートである。

【図 6】 本発明の一実施例の情報の例である。

【図 7】 本発明の一実施例のシソーラスの例である。

【図 8】 本発明の一実施例の付加情報データベースの概念図である。

【図 9】 本発明の一実施例の情報探索の例を説明するための図である。

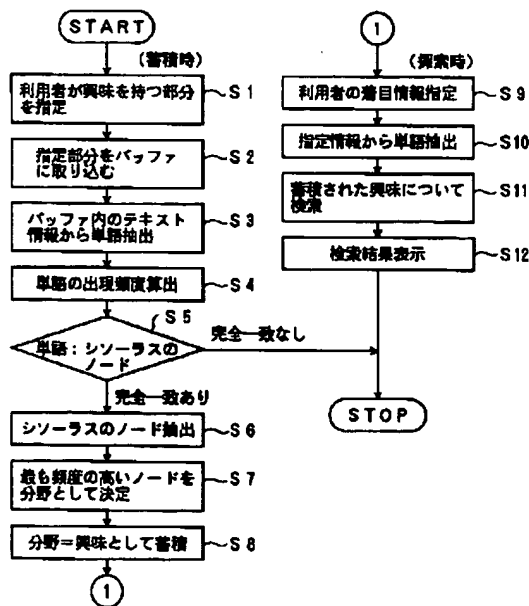
## 【符号の説明】

10 検索モジュール  
 11 領域選択部  
 12 情報検索部  
 13 情報表示部  
 20 蓄積モジュール

21 領域選択部  
 22 情報分類部  
 23 情報蓄積部  
 30 情報  
 40 付加情報  
 50 付加情報記憶手段、付加情報データベース  
 60 シソーラス、シソーラス格納部  
 110 参照情報指定手段  
 120 情報指定手段  
 210 指定部分格納手段  
 220 記憶手段  
 230 単語抽出手段  
 240 単語出現頻度計算手段  
 250 ノード抽出手段  
 260 分野決定手段  
 270 興味情報格納手段  
 280 指定単語抽出手段  
 290 興味検索手段  
 295 結果表示手段

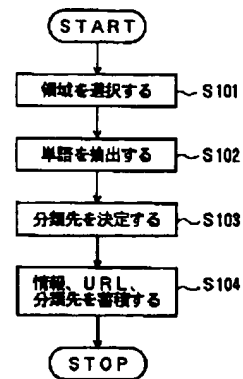
【図 1】

本発明の原理を説明するための図



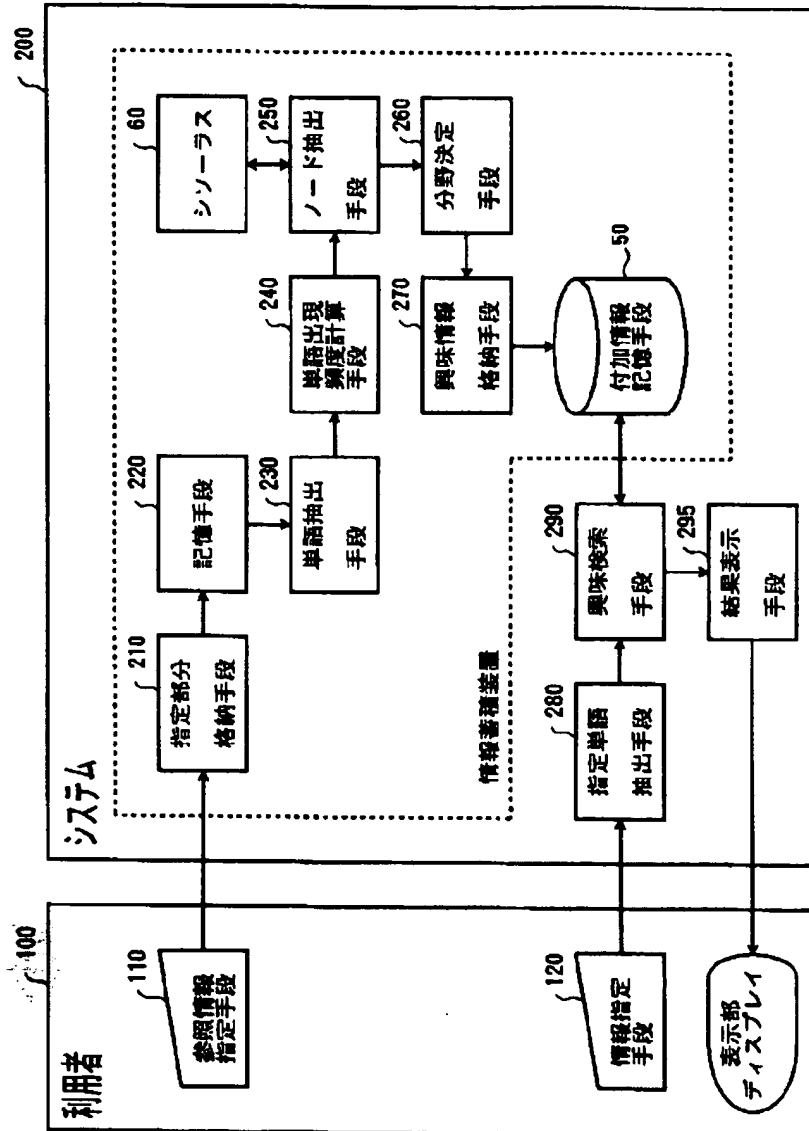
【図 4】

本発明の蓄積モジュールの動作を示すフローチャート



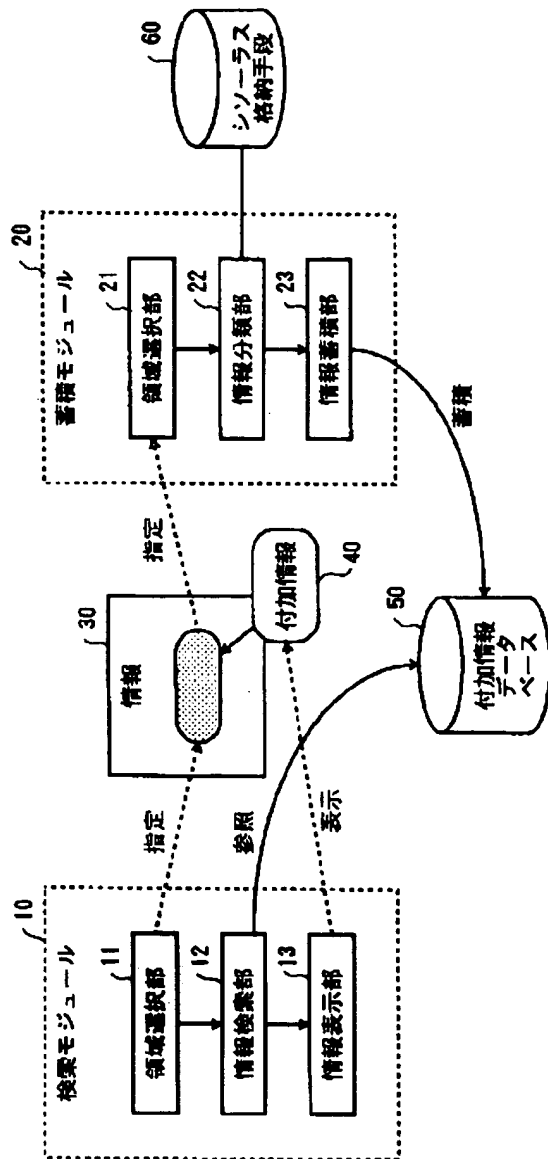
【図 2】

## 本発明の原理構成図



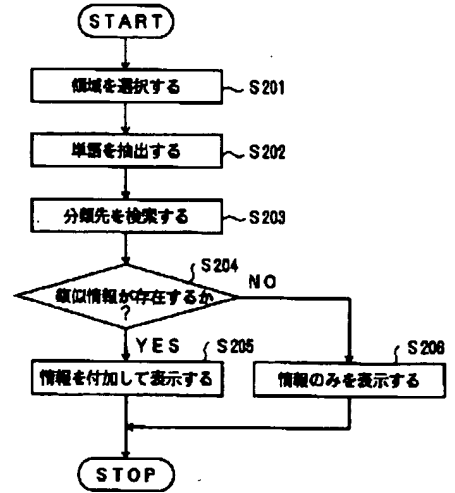
【図3】

本発明の情報推薦システムの構成図



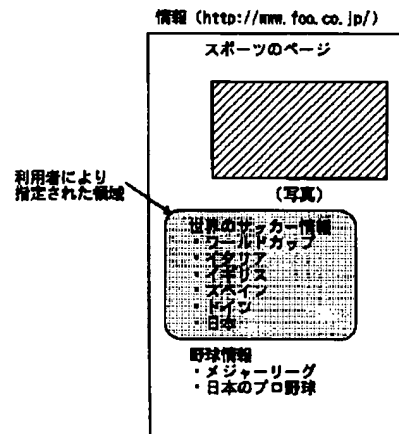
【図5】

本発明の検索モジュールの動作を示すフローチャート



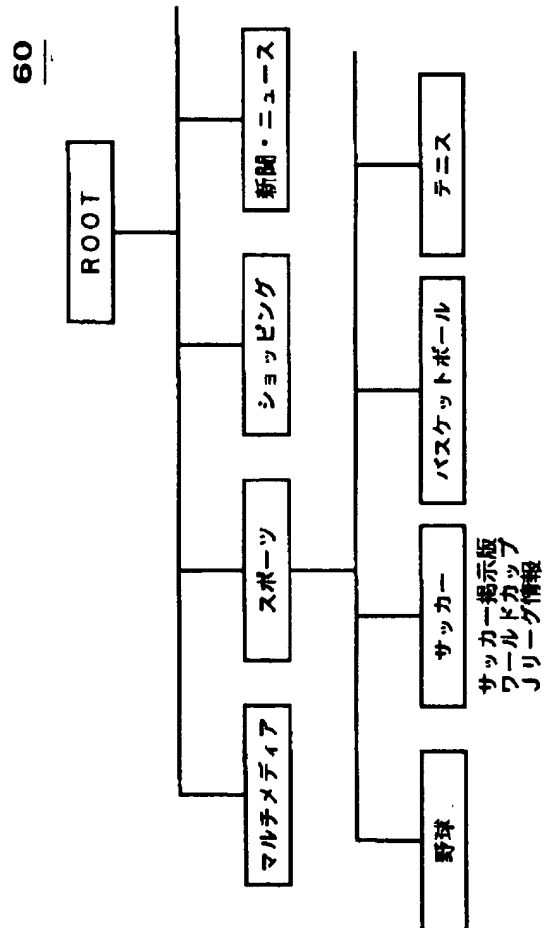
【図6】

本発明の一実施例の情報の例



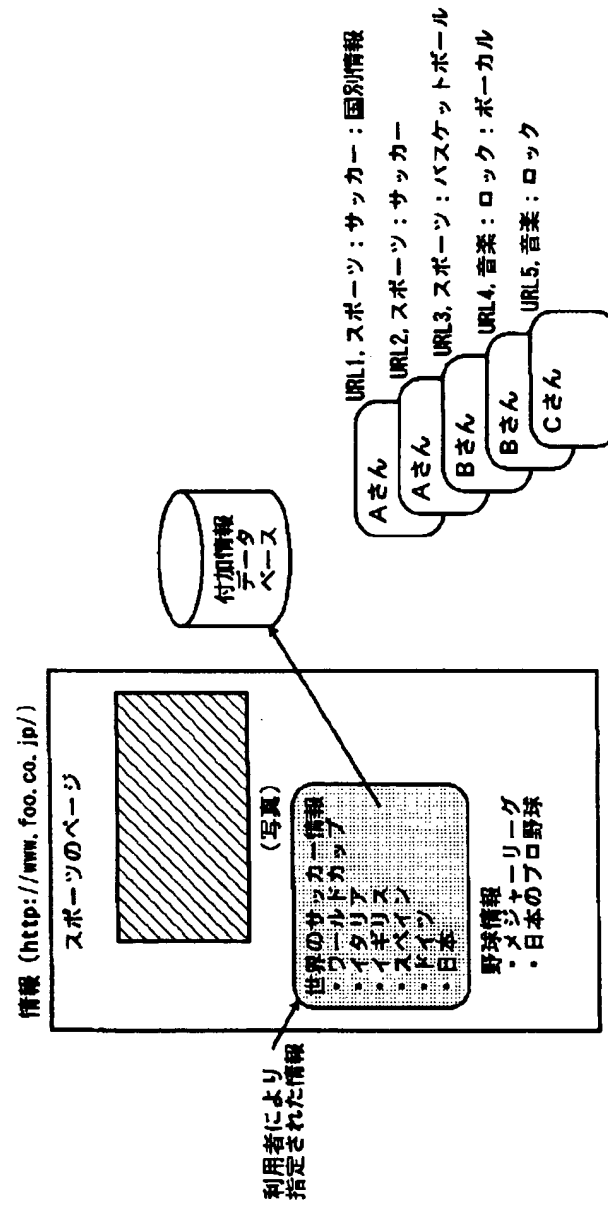
【図 7】

本発明の一実施例のシソーラスの例



【図8】

## 本発明の一実施例の付加情報データベースの概念図



【図9】

本発明の一実施例の情報探索の例を説明するための図

